National Center for Health Statistics





State-level Estimation Using the National Health Interview Survey (NHIS)

Stephen Blumberg

Director, Division of Health Interview Statistics

Meeting of the NCHS Board of Scientific Counselors October 22, 2021

Recent State Estimates from the **National Health Interview Survey**

National Health Statistics Reports

Number 163 August 6, 2021

Geographic Variation in Health Insurance Coverage: United States, 2019

by Robin A. Cohen, Ph.D., Emily P. Terlizzi, M.P.H., Amy E. Cha, Ph.D., M.P.H. Michael E. Martinez, M.P.H., M.H.S.A., Van L. Parsons, Ph.D., Rong Wei, Ph.D., and Yulei He, Ph.D.

Abstract

Objectives-This report presents state, regional, and national estimates of the percentage of persons who were uninsured, had private health insurance coverage, and had public health insurance coverage at the time of the interview.

Mathods... Data from the 2019 National Health Interview Survey were used to estimate health insurance coverage. Estimates were categorized by age group, state Medicaid expansion status, urbanization level, expanded regions, and state. Estimates by state Medicaid expansion status, urbanization level, and expanded regions were based on data from all 50 states and the District of Columbia. State estimates are shown for 32 states and the District of Columbia

Results-In 2019, among persons under age 65, 12.0% were uninsured, 64.3% had private coverage, and 25.9% had public coverage at the time of the interview. Among adults aged 18-64 (working-age adults), the percent uninsured ranged from 12.4% for those living in large fringe (suburban) metropolitan counties to 17.5% for those living in nonmetropolitan counties. Working-age adults living in non-Medicaid expansion states (20.8%) were about twice as likely to be uninsured compared with those living in Medicaid expansion states (10.9%). Similar patterns were observed among children aged 0-17 years. The percentage of working-age adults who were uninsured was significantly higher than the national average (14.5%) in Florida (20.6%), Georgia (22.3%), Oklahoma (25.6%), and Texas (30.5%), and significantly lower than the national average in California (11.5%). Minnesota (6.9%). New York (7.4%). Ohio (10.8%), Pennsylvania (9.8%), and Wisconsin (7.7%). The percentage of people under age 65 who were uninsured was lowest in the New England region (4.6%).

Keywords: uninsured * private * public * state level * National Health Interview

Introduction

Health insurance coverage in the United States is a key measure of health care access (1-3). Previous receased based on national conceve. has found geographic variation in insurance coverage in the United States by urbanization level, state Medicaid expansion status, region, and state (4-6). Population estimates of health insurance coverage at the state level are necessary for the development and assessment of federal and state health care coverage programs and policies (7-9). A recent study found that more than 4 million persons would gain coverage if the remaining non-Medicaid expansion states would fully implement a Medicaid expansion under the provisions of the Affordable Care Act (ACA) (10.11).

This report examines geographic variation in health insurance coverage in the United States in 2019. Estimates of the percentage of persons who were uninsured, had private coverage, and had public coverage at the time of the interview are presented by urbanization level, state Medicaid expansion status, expanded regions, and selected states. The primary focus of this report will be on persons under age 65, because nearly



Health insurance coverage for 32 states

Household telephone service estimates for 50 states





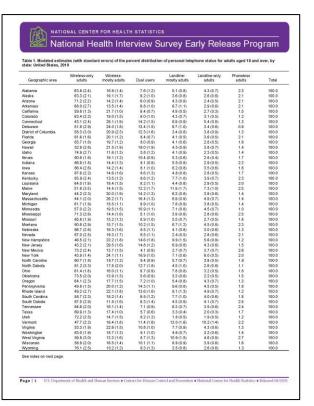


U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES Centers for Disease Control and Prevention National Center for Health Statistics

CDC

NCHS reports can be downloaded from: https://www.cdc.gov/hchs/broducts/index.htm





A Brief History of State Estimates from the **National Health Interview Survey**

DATA EVALUATION AND METHODS RESEARCH

Number 75

Synthetic Estimation of State Health Characteristics Based on the **Health Interview Survey**

This report discusses the various methods that have been proposed or used for obtaining estimates of health characteristics for local areas. Particular emphasis is given to discussion and evaluation of synthetic estimation procedures developed originally at the National Center for Health Statistics for purposes of estimating levels of health characteristics obtained from the Health Interview Survey for each State and the District of Columbia.

DHEW Publication No. (PHS) 78-1349

U.S. DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE Public Health Service National Center for Health Statistics Hyattsville, Md. October 1977



1960s -1980s:

Ratio-adjusted synthetic estimates

1990s: Bayesian methods using hierarchical models

Small Area Inference for Binary Variables in the National Health Interview Survey

Donald MALEC, J. SEDRANSK, Christopher L. MORIARITY, and Felicia B. LECLERE

The National Health Interview Survey is designed to produce precise estimates of finite population parameters for the entire United States but not for small geographical areas or subpopulations. Our investigation concerns estimates of proportions such as the probability of at least one visit to a doctor within the past 12 months. To include all sources of variation in the model, we carry out a Bayesian hierarchical analysis for the desired finite population quantities. First, for each cluster (county) a separate logistic regression relates the individual's probability of a doctor visit to his or her characteristics. Second, a multivariate linear regression links cluster regression parameters to covariates measured at the cluster level.

We describe the numerical methods needed to obtain the desired posterior moments. Then we compare estimates produces

using the exact numerical method with approximations. Finally, we compare the hierarchical Bayes estimates to empirical Bayes estimates and to standard methods, that is, synthetic estimates and estimates obtained from a conventional randomization-based approach. We use a cross-validation exercise to assess the quality of model fit. We also summarize the results of a senarate study f the binary indicator of partial work limitation. Because we know the value of this variable for each respondent to the 1990 Census long form, we can compare estimates corresponding to alternative methods and models with very accurate estimates of

KEY WORDS: Bayesian predictive inference; Cluster sampling; Cross-validation; Empirical Bayes; Hierarchical model; Logistic regression; Synthetic estimates

1. INTRODUCTION

The National Health Interview Survey (NHIS) is a multistage, personal interview sample survey conducted annually for the National Center for Health Statistics (NCHS) to provide information about the health and health care utilization of the civilian, noninstitutionalized population in the United States. Basic health questions are asked in all selected households (e.g., perceived health status, visits to a physician). Persons in subsampled households are also in these rates in 78 small areas. Because federal funding to asked more detailed questions about chronic and acute conditions (e.g., hypertension, bronchitis). Special health topics are investigated using a subsample of individuals. The hasic core questions and questions about chronic and acute diseases are covered annually, whereas questions on special topics are asked only periodically.

The 1985-1994 NHIS sample has about 200 primary sampling units (PSUs), selected from a stratified population of size 1,983. Each PSU consists essentially of a single a thorough, excellent review, and an older article by Purcell county or a group of contiguous counties. Within each sampled PSU, groups of households are aggregated into areal segments and sampled. Each year there is a new sample with small area estimation projects in the U.S. government see approximately 50,000 households and 120,000 individuals. Massey Moore Parsons and Tadros (1989) provided additional details about the design of the NHIS.

The NHIS emphasizes national estimates, but there is also a need for estimates for small subpopulations or geo-

Donald Malec is Mathematical Statistician, Christopher L. Moriarity is 1979). Thus we investigate alternative estimators based on Mathematical Statistician, and Felicia B. LeClere is Health Statistician, National Center for Health Statistics, Centers for Disease Control and Prevention, Hyattsville, MD 20782, J. Sedransk is Professor, Department of Statistics. Case Western Reserve University, Cleveland, OH 44106. The authors are grateful to Linda Tompkins, Brian Morrow, John Edmonds, and Gary Gensler for providing programming support; to Richard Evans, Jiming Jiang, Partha Lahiri, and Balgobin Nandram for their comments; to Althelia Harris, Karen Whitaker, Carol Larson, and Ladd Sheets for their areful typing of the manuscript; and to the editor, Diane Lambert, for her patience and helpful suggestions. Sofransk is grateful to the Office of Research and Methodology, National Center for Health Statistics, for its support of his research on this project.

graphical areas. For example, Lieu, Newacheck, and Mc-Manus (1993) used data from the 1988 NHIS child health supplement to compare access to health care and use of doctors for 10-17-year-old blacks, whites, and Hispanics. Sometimes estimates are required for small subnopulations within small geographical areas. For example, the National Immunization Survey provides estimates of immunization rates for children ages 19-35 months and monitors changes on the immunization coverage rates, precise point estimates and reliable measures of variability are needed.

Our objective is to provide appropriate estimates from the NHIS for the 50 states and the District of Columbia and for salient subpopulations within these 51 areas. Many methods exist for inference about small geographical areas and dotailed descriptions of some methods. For a recent review of Schaible (1995). Unfortunately, randomization-based estimates (also known as design-based or direct estimates), are not sufficiently precise for many states and subpopulations under the current design (with sampling roughly proportional to PSU population size). A commonly used method, synthetic estimation, has also been shown to be unsatisfactory (Särndal 1984; Schaible, Brock, Casady, and Schnack realistic models, and compare these with the standard esti-

Most of the variables in the NHIS are binary, and we use hierarchical models similar to those used by Dempster and Tomberlin (1980) and Wong and Mason (1985). We use

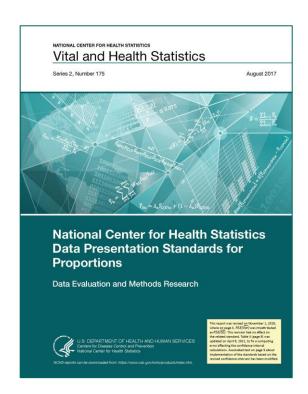
© 1997 American Statistical Association rnal of the American Statistical Association September 1997, Vol. 92, No. 439, Applications and Case S

Milestones in State-level Health Insurance Estimates

- 2004: Direct "Early Release" estimates for 10 states
 - Based on data from one year
 - Expanded to 20 states from 2005-2010
- 2008: Direct estimates for 41 states
 - Based on pooled data from three years (2004-2006)
 - Estimates reported for states with 1000+ persons in pooled sample
 - Smoothed variance estimates using design effect estimated from the 10 most populous states
- 2014 and 2015: Direct estimates for all 50 states
 - Augmented sample size allowed for single-year estimates for all states
 - Otherwise, the number of states with estimates varied by annual survey sample size from 2011-2018, ranging from 17 to 45

Recent Changes Impacting Ability to Produce Direct State Estimates

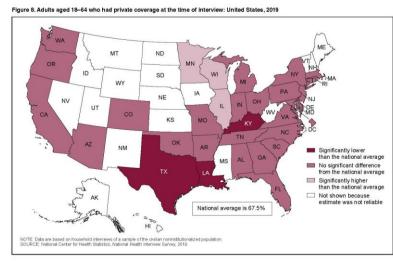
- 2019: NHIS questionnaire redesign
 - Person counts reduced even though household counts remained the same
- 2017: New NCHS data presentation standards
 - Added a standard regarding minimum degrees of freedom
- 2016: New NHIS sample design
 - Stratified, clustered sample with 534 clusters nationally
 - Clusters do not cross state lines
 - 18 states have fewer than 8 clusters
 - Clusters remain unchanged over 10-year span of the design



Relaxing Requirement that Clusters Remain Unchanged

- NHIS sample will be redesigned for 2025
- One goal is to improve ability to produce state estimates
- In 2022-2024, Census Bureau will examine impact of rotating clusters on costs and workforce retention
 - Pilot test in 6 states, 4 of them selected from those states with fewer than 8 clusters



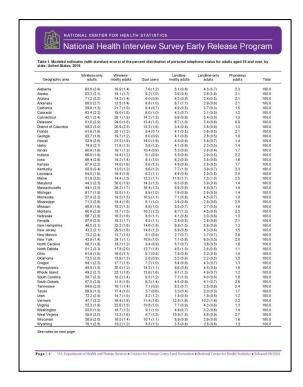


States with estimates not shown are the 18 states with fewer than 8 clusters each

Figure Source: Cohen RA, Terlizzi EP, Cha AE, Martinez ME, Parsons VL, Wei R, He Y. Geographic variation in health insurance coverage: United States, 2019. National Health Statistics Reports; no 163. 2021.

Model-Based Estimates

- NCHS has published state-level telephone service estimates annually since 2011
 - Estimates of wireless-only, landline-only, and dual users for all 50 states and DC
 - Derived using a weighted combination of a direct NHIS estimate and a regressionsynthetic estimator that uses data from NHIS and ACS
 - Developed and implemented every year by Nada Ganesh (NORC)



Small-Area Estimation Project: A Feasibility Study

Recently launched with the Census Bureau's Small Area Modeling and Development Branch

Project Goals

- 1. Develop models that will produce annual state-level small area estimates
- 2. If possible, develop models that will produce annual state-level small area estimates with demographic detail
- 3. Evaluate the performance of models that use publicly available auxiliary data compared with those that use restricted data
- 4. If possible, seek a generalizable framework for future small area estimates

• Initially producing state-level estimates for four measures:

Foregone medical care due to cost; asthma; asthma episodes;
 and self-reported worry, nervousness, or anxiety

Potential Topics for BSC Discussion

- How best to achieve and evaluate state estimates
- Value of state estimates from NHIS
- Value of subgroup estimates within states, defined by geography or sociodemographic groups